Attn: Mr. Behrooz Senfi, Art Unit 2613
Commissioner of Patents and Trademarks
Washington, D.C. 20231

June 12, 2006

Dear Mr. Senfi:

This communication is attached to the accompanying Request for Continuation and is in regard to Patent Application Serial No. 10/006,444 entitled "Optimizations for Live Event, Real-time, 3D Object Tracking." Please find enclosed my check number ~~1436~~ 1058 covering the $395 RCE fee for small entities.

In response to your Office Action dated January 11, 2006 and my office visit and interview on May 03, 2006, please cancel Claims 34-49 and add attached new Claims 50-96. Also please find my enclosed check number 1059 ~~1437~~ covering the two month extension cost of $225 for responding to the office action within five months of its mailing date.

With respect to these new claims, there are four independent claims 50, 61, 73 and 84. Of these, claims 61 and 84 are fully supported in the specification of continued application number 09/197,219, filed on Nov. 20, 1998, now Patent No. 6,567,116, and are therefore not new subject matter. I am requesting that since these claims have proper antecedent basis within continued application 09/197,219, that they be considered for novelty based upon the filing date of Nov. 20, 1998. Attached please find Addendum A, "Subject Matter Support for Claims 61 and 84," that provides detailed references from application 09/197,219 that show proper antecedent basis.

With respect to my office visit and interview on May 03, 2006, please note that the following changes were made to the proposed new Claims now 50-96:

1. The word "operator" was replaced by "user";
2. The words "fixed cameras" were replaced by "stationary cameras";
3. The use of "(X, Y)" and "(X, Y, Z)" was replaced with "X, Y" and "X, Y, Z";
4. The word "location" was replaced by "coordinates", and
5. The words "dimensional characteristics" were expanded to "dimensional characteristics of at least the size of each object."

With respect to the Office Action, the following list outlines the sections of my response:

1. A summary of major differences between my teachings and that of Paff's;
2. Detailed response to Point 3 of the Office Action;
3. A side-by-side comparison of the changes to my claims, and
4. Arguments for the allowance of the revised claims.

*(1) A summary of the major differences between my teaching and that of Paff's:*

My patent teaches a system for first automatically tracking one or more objects in a predefined area using video analysis captured from a first set of fixed cameras and then second automatically adjusting one or more movable second cameras to follow the one or more tracked objects. Paff teaches a system for first manually adjusting a movable Master Camera's view to follow a single object or single group of objects, second automatically tracking the changes to the Master Camera's view, and then third automatically adjusting one or more movable Slave Cameras to follow the tracked Master Camera's view.

The following table compares our respective apparatus:

| | Part Description | Paff | Aman |
|---|---|---|---|
| Part A | Apparatus necessary to determine center-of-view(s). | *One* pan, tilt, focus and zoom *adjustable* Master Camera that provides a *partial-view* of the predefined area and outputs a continuous First Video Stream;<br><br>A Monitor for viewing the continuously output First Video Stream;<br><br>A Joy Stick for allowing an operator to continuously control the Master Camera, where the Joy Stick outputs a continuous stream of Control Signals representative of the operators directives that is first used to adjust the Master Camera's view;<br><br>A human operator to control the Joy Stick and thereby move the Master Camera's center-of-view to follow *one* object or *one* group of objects, and<br><br>A Controller for continuously *analyzing the stream of Joy Stick Control Signals* to determine *one* continuous set of (X, Y) coordinates for the *one* "selected position" representing the center-of-view of the Master Camera. | A First *set* of *Fixed* Cameras combined to create a contiguous *full-view* of the predefined area and outputting a continuous First Video Stream, and<br><br>A First algorithm for continuously *analyzing the First Video Stream* to determine *one* continuous set of (X, Y) coordinates for *each* of *multiple* objects, where each object represents a distinct center-of-view. |

| | | | |
|---|---|---|---|
| Part B | Apparatus for creating additional perspective views based upon determined center-of-view(s). | A Second Set of adjustable Slave Cameras that continuously change their individual center-of-views to match the *one* center-of-view determined by Part A that output a Second Video Stream for operator viewing. | A Second Set of adjustable Cameras that continuously change their individual center-of-views to create best views of multiple objects (e.g. depending upon occlusions) based upon the *multiple* center-of-views determined by Part A that outputs a Second Video Stream. |
| Part C | Apparatus for determining additional object information from additional perspective views. | | A Second algorithm for continuously analyzing the Second Video Stream to determine *one* continuous set of (X, Y, Z) coordinates for *multiple* recognizable features on the *multiple* objects. |

The following table compares our methods:

| | Step Description | Paff | Aman |
|---|---|---|---|
| Step 1 | Determine a center-of-view. | *Operator manually* looks at a *partial-view* of the predefined area using *one adjustable* Master Camera that is user controlled by a Joy Stick that outputs *Control Signals* for adjusting the Master Camera's center-of-view, and<br><br>System automatically analyzes *the Control Signals* to determine the (X, Y) coordinates of the *one* center-of-view of the Master Camera. | *System automatically* "looks at" the *full-view* of the predefined area using a First *Set* of *Fixed* Cameras that outputs a First *Stream of Video*, and<br><br>System automatically analyses *the First Stream of Video* to determine the (X, Y) coordinates for *each* object. |
| Step 2 | Adjust view of surveillance / video cameras to match center-of-view. | System uses the automatically determined (X, Y) location of the *one* Master Camera center-of-view to automatically adjust and the viewing angles of each Slave Camera to overlap the *one* center-of-view. | System uses the automatically determined (X, Y) locations of the *multiple* tracked objects to automatically adjust the viewing angles of each Second Camera to best overlap each of the *multiple* objects, where each Second Camera outputs to a Second Stream of Video. |

| | | | |
|---|---|---|---|
| Step 3 | Create additional feature location information. | | System automatically analyses the *Second Stream of Video* to automatically determine the (X, Y, Z) coordinates for *each* recognizable feature on *each multiple* object. |

***(2) Detailed response to Point 3 of the Office Action:***

Point three of the OA begins with a direct comparison between my claim 34 and the teaching's of Paff's as follows:

> "Regarding claim 34, Paff discloses, an automated system for tracking the movement of multiple objects within a predefined area (figs. 2-3, col. 8, lines 6 – 10) comprising: a first set of cameras forming a fixed area tracking matrix for first detecting the motion of each object in (X, Y) space (figs. 2 – 3, camera MD) and first algorithm operated on a computer system responsive to the fixed area tracking matrix for determining the (X, Y) location of each object, (figs. 5 and 6, controller 10, includes a mathematical processor to perform mathematical calculation with respect to position of the subject) and a second set of cameras forming a movable volume tracking matrix responsive to the determined (X, Y) locations from the first algorithm for controllably detecting the motion of each object in (X, Y, Z) space (fig. 6, slave cameras, col. 4, lines 50 – 60) and a second algorithm operated on the computer system responsive to the second set of cameras for determining the (X, Y, Z) dimensional characteristics of each object, (fig. 6, each slave camera has the same component configuration as the master camera, each has a controller 10, which includes a mathematical processor to perform mathematical calculation based on the information received from the master camera)."

The following outlined arguments breakdown this initial assessment into clauses 1 – 14. In detailed response to each clause, explanations are respectfully provided as to why the interpretations used in the comparison draw incorrect conclusions about my teachings and / or Paff's. Specifically:

> "Regarding claim 34, Paff discloses,"

1. "an *automated* system"
   a. Two common definitions for "automation" are:
      i. "Starting or functioning by itself, i.e. started, operated, or regulated by a process or mechanism *without human intervention*", and
      ii. "*Machine operating without human intervention*, i.e. a machine, for example, a washing machine, that controls its own operating process."

b. Any "automated" system by common understanding does not require continuous operator intervention in order to function.

c. Paff's teachings clearly require continuous operator intervention, specifically Col. 4, lines 23-35 state:

"*An operator*, located at the monitoring station 11, views a primary monitor 11A which displays the video output of the master camera MD. The operator can also view a number of secondary monitors 11C, 11D, etc. which display the video outputs of one or more of the slave cameras SD1-SD5. *By operating a joy stick* 11B, or a similar device, *the operator can transmit* messages to the controller 10 of the master camera MD *requesting* that the pan angle and the tilt angle of the image and lens assembly 17 of *the camera be changed.* The controller 10, in response to the received request, drives the pan motor 13 and/or tilt motor 14 in the desired direction." (italics added)

d. Hence, at least this portion of Paff's teaching is clearly not automated. Furthermore, this is a necessary firs step of my teachings – i.e. to "determine the (X, Y) location of the multiple objects" of interest. Nowhere in either my present application or in any of its included references is an operator required for any continuous functions, i.e. for anything more than initial setup type work. Paff has not automated this necessary first step whereas my teachings do. My teachings show that the current locations of the objects are determined using image analysis algorithms to continuously search through the streams of images being collected by the first set of fixed cameras. Paff has no such equivalent.

e. After requiring an operator to manually look at a monitor and operate a joystick to control the pan and tilt angle of a master camera as it follows a single object, Paff *then* teaches an *automated* system for determining where the operator is "looking" and then without further operator intervention directing slave cameras to also follow the operator selected center-of-view. At best, Paff can be considered a "semi-automated" system and this difference between our teachings alone is extremely significant. For instance, a system such as Paff's being used to video a youth hockey game would absolutely require an individual to actively and continuously video the game action from one viewpoint after which Paff's apparatus would automatically collect additional video from alternate additional viewpoints.

    i. This single requirement of an operator makes Paff's device unacceptable and unworkable for situations where an operator is not available and clearly differentiates his teaching from my specification and claims.

2. "for tracking the movement of *multiple* objects within a predefined area"

    a. It is perhaps initially a subtle difference in teachings, but I respectfully submit that Paff at most teaches the tracking of a single object per single operator / master camera per any given moment. Furthermore, this subtle difference has great significance as it renders Paff's invention unworkable as a means for simultaneously collecting the ever-changing (X, Y) locations of more than one center-of-view / object. This is especially true for larger predefined areas where the multiple objects are significantly spread out and will not "fit" into the view of

the master camera, even when fully zoomed out. And even if all the multiple objects where to fit into the master camera's view, they might all be moving in different directions with therefore different current (X, Y) coordinates. Paff's system would require the operator to manually pick a center-of-view for the "single group" of multiple objects after which his system would then track the (X, Y) coordinates of that "single group." Hence, for Paff "*a single group* of multiple sub-objects" really equals a "*single object.*"

i. Furthermore, the operator chosen center of a moving group of objects may in fact not represent the location of any of the actual objects being viewed. This would easily be the case where two people were running in parallel separated by even some small distance, such as 3 feet. In this case, the operator might be fixing their view on the area in between the moving people so as to center the Master Camera's view.

b. The best and most obvious proof that Paff is only tracking a "single" object / group / entity can be reasoned based upon the following excerpts from Paff's patent:

Col 4, lines 42-49: "As the master camera MD is moved, the controller 10 periodically calculates, based on the aforesaid stored current pan and tilt angles and the stored master camera's height above and its projected coordinate position in the plane 2, *the coordinates* of the position (*the selected position*) in the plane intersected by the optical axis 17A of the camera. The calculated coordinates are then broadcast to all slave cameras SD1-SD5." (italics added)

i. My emphasis on *"the coordinates"* ... *"the selected position"* is meant to point out the singular nature of the "object tracking data" generated by Paff's system. His system produces a times series of (X, Y) coordinates that follow a *single, "selected position" not plural "positions"*.

ii. My system clearly generates a time series of (X, Y) coordinates for *each and every* object moving within the predefined area. This is why my teachings truly describe how to track multiple objects whereas Paff's tracks only a single object per operator / master camera – which is also his claimed system (i.e. he does not claim "multiple master cameras" for "tracking multiple objects" but rather a single master camera that tracks a single position that is automatically followed by multiple slave cameras.)

c. Furthermore, in Col 8, lines 6-13, Paff states:

"In this manner, if the zooming state of the master camera is set to wide angle so that *a large group* of subjects can be tracked through the premises, the slave cameras will also be set to a wide angle position. Conversely, if the zooming state of the master camera is set tight so as to track *a single individual* through the premises, the slave cameras similarly will have a tight zoom setting." (italics added)

i. Notice that Paff can either track "a (single) large group" or "a single individual". Just because the "(single) large group" contains multiple

"subjects" does not mean that Paff's system is generating a centered (X, Y) coordinate for each subject in the group.

ii. My teachings clearly show how to first create a separate stream of (X, Y) coordinates for each and every object of multiple objects within the predefined area. Afterwards, one or more second movable cameras (i.e. "slave cameras" to Paff) can be directed to follow *any* of these individually tracked separate objects.

1. This is a critical difference between our teachings and is specifically illustrated and described by Fig. 16b of my application.

3. "(figs. 2-3, col. 8, lines 6 – 10)"

a. The OA refers to the following lines of Paff's specification as indicative of his teaching a system that tracks the movement of multiple objects with a predefined area, specifically, Col. 8, lines 6-10 read:

"In this manner, if the zooming state of the master camera is set to wide angle so that a large group of subjects can be tracked through the premises, the slave cameras will also be set to a wide angle position."

b. As previously stated, to Paff the large group becomes a single object for which a continuous stream of (X, Y) "selected position" coordinates are generated. Just because the large group is made up of multiple subjects does not mean Paff is tracking these same multiple subjects. In fact the operator is intuitively selecting a best-fit middle of the group as their "selected position" where this middle of the group may not even represent the location of any single subject / object.

4. "comprising: a first set of cameras forming a fixed area tracking matrix"

a. While Paff uses a "first camera" in the form of his MD master camera, this is only a single camera and even when zoomed out, for larger areas the single master camera is insufficient to view the entire predefined area.

b. In Col 7, lines 34-37 Paff states:

"It should be noted that any of the cameras in the system 1 can be switched to take the role of the master camera. The operator at the station 11 can effect this change."

c. At first glance, Paff might seem to be using multiple cameras to continuously view the entire predefined area except for the following considerations:

i. None of these master or slave cameras are fixed in their view and therefore cannot even collectively guarantee continuous coverage of the tracking area;

ii. The slave cameras by Paff's definition are continuously moved to follow the same view as the master camera. This implicitly means that they are

all continuously excluding any portion of the predefined tracking area not currently "selected" by the master camera;

iii. While the operator can switch between any camera "in the set" of master and slave cameras at any time, only *one* of the cameras (i.e. the chosen master camera) is doing the object tracking at any single instant.

d. Hence, Paff comparatively teaches "a first set of cameras forming a movable area tracking view from which any of the cameras may be chosen by an operator to have that camera's manually adjusted optical axis be used for the automatic determination of the single 'selected position' / center-of-view."

5. "for first detecting the motion of each object in (X, Y) space"

   a. As previously stated:

      i. Paff at most only determines the (X, Y) motion of a single object, not "each object" of multiple possible objects.

      ii. Paff determines the (X, Y) coordinate of the "selected position" as viewed through a Master Camera, which technically is not that same as the location of even a single object, let alone multiple objects.

         1. It is impractical to expect a human operator to keep a camera view centered on an athlete during a competition. This reality precludes accurate object position tracking especially for high acceleration / deceleration scenarios like a sporting event.

6. "(figs. 2 – 3, camera MD)"

   a. My teachings do not specify the use of a master camera, wherein each and every object to be tracked must be viewable through the single master camera.

   b. Furthermore, as previously stated, Paff's Master Camera is so defined because it is the camera being controlled via Joy Stick by an Operator (otherwise, it would be a Slave Camera.) Hence, for best comparison, Paff's "camera MD" would be better labeled "camera MD with operator," which is clearly different from my teachings.

7. "and first algorithm operated on a computer system responsive to the fixed area tracking matrix"

   a. Paff's approximation of my "first algorithm" is his "controller 10" as stated in Col 4, lines 36-49 as follows:

      "The controller 10, in response to the received request, drives the pan motor 13 and/or the tilt motor 14 in the desired direction. During this movement, the feedback outputs 13A and 14A of the motors provide the controller 10 with and exact value for the current pan angle of the master camera MD and an exact value for the tilt angle of the camera. These values are stored by the controller 10 in its memory 12. As the master camera MD is moved, *the controller 10 periodically calculates, based on the aforesaid stored current pan and tilt angles*

*and the stored master camera's height above and its projected coordinate position in the plane 2, the coordinates of the position (the selected position)* in the plane intersected by the optical axis 17A of the camera. The calculated coordinates are then broadcast to all slave cameras SD1-SD5."

b. As specified by Paff, the controller is *actually* responsive to "*the aforesaid stored current pan and tilt angles and the stored master camera's height above and its projected coordinate position in the plane 2*" (i.e. the motion control signals) and *not* the *image data* output by even the master camera, let alone a "fixed area tracking matrix" (which Paff does not have.)

    i. My patent teaches that the first algorithm operates on the *image data* created by a first set of fixed cameras. This data is clearly different then Paff's motion control signals created by Joy Stick movements and has significant benefits, at least including:

        1. the ability to determine to location of multiple objects in the predefined area, not just the assumed center-of-view of a single master camera;

        2. the ability to determine the size of each tracked object;

        3. the ability to determine the orientation of each tracked object, and

        4. the ability to either generally classify and / or even identify the tracked objects.

    ii. All of these additional significant benefits *require* image data and image analysis.

        1. Not only are these benefits not even discussed by Paff, his specification cannot reasonably be construed as capable of creating any of this important information.

8. "for determining the (X, Y) location of each object,"

a. As previously stated, Paff is not technically determining the (X, Y) location of the object being "sighted," but rather he is determining the location of the master camera's center-of-view. This is a subtle but important distinction for the following reasons:

    i. Since image analysis is not being used, nor any form of beacon tracking on the object, the master camera center-of-view is necessarily determined / adjusted by an operator. This reliance on operator decisions with the inherent opportunity for human error is one of the key motivating factors for automation.

    ii. In many situations, the current (X, Y) location of the object will not be equivalent to the ideal center-of-view for the slave cameras. Hence, with Paff's system the slave cameras follow a simple rule that they should always be centered on the same view as the master camera, regardless of other potential circumstances, such as object occlusion (or blocking).

1. In my teachings, I discuss the need to first determine where each object is in relative (X, Y) space so that this location can at least be compared to:

    a. The positions of all second (Paff's slave) cameras, and;

        i. (Paff does this.)

    b. The positions of any and all additional objects also within the predefined area.

        i. (Paff does not do this.)

2. These two comparisons are both extremely important for choosing the proper orientation of each second (slave) camera, since any additional objects may ultimately be blocking a particular second (slave) camera's current view, thereby necessitating a more sophisticated center-of-view selection strategy.

iii. Paff's method for triangulating to the assumed object's position based upon the master camera's current pan and tilt angle as well as its height above the tracking surface can create substantial errors in the (X, Y) location of a current object, as substantiated below:

1. In Col. 3, line 63 through Col. 4, line 5, Paff teaches:

    "The reference plane 2 defines the points or positions in the premises 1 which can be selected by the master camera MD. The master camera MD selects a position by causing the optical axis 17a of its image and camera assembly 17 to be focused on and intersect the position. By appropriately selecting the height of this plane relative to the height of the average human, the master camera MD, when directed at a position, will be able to view an entire human subject if located at the selected position. A typical height for the reference plane 2 might be four feet from the premises floor."

2. This strategy for assuming an "intersect" point (e.g. 4 feet) relative to the "premises floor" is significantly inadequate for sporting events where the objects being tracking will experience a wide variation in this "typical height" value.

    a. Hence, Paff's strategy is only reasonable for a consistently sized object maintaining a consistent distance from the floor, such as a walking human and is unworkable for at least the highly variable movement of a game object such as a puck, football or basketball and also for game athlete that may alternately fall, leap, stand and bend over.

    b. Under such circumstances as trying to follow a moving basketball, football or hockey puck, Paff's approach will

incorrectly assign the center-of-view and therefore will also incorrectly adjust the views of all slave cameras.

3. Paff assumes that the master camera will be at an angle sufficiently above the assumed (4 ft) height of the target (human) object that the line of it's optical axis can and will only intersect the target over a small range of total distance from the master camera.

    a. This assumption will cause Paff's system to incorrectly assign the center-of-view when the master camera is closer to the ground surface and looking substantially straight out over a longer distance towards the object of interest, such as an end-zone camera viewing the entire field in which case players over a wide range of distance can all be misinterpreted as the "selected position."

9. "(figs. 5 and 6, controller 10, includes a mathematical processor to perform mathematical calculation with respect to position of the subject)"

    a. The mathematical calculations taught by Paff are substantially different than the image analysis techniques discussed in my application and cannot be construed to teach my underlying technique.

    b. Specifically, the dynamic portion of Paff's calculations is based upon the ever changing variables of the master cameras current tilt and pan angles. These angle measurements are entirely different from the dynamic video images provided for analysis to my first algorithm.

10. "and a second set of cameras forming a movable volume tracking matrix responsive to the determined (X, Y) locations from the first algorithm"

    a. It is important to note that the slave cameras in Paff's patent are responsive to a single stream of (X, Y) locations (*assumed* to be the center of a single object, or a single group of objects,) whereas my second set of cameras is responsive to multiple streams of (X, Y) locations (which are the actual detected centers of each object.)

    b. In the discussion of my Fig. 16b, my specification states:

    "Referring now to **Fig. 16b**, there is shown an example matrix of four FOV's **120v** created by area tracking cameras **124**. Within this combined grid, *several players* having top surfaces such as **110x** and **111x** move freely about. In this particular example, *four movable cameras* **140-a, 140-b**, **140-c** and **140-d** *are tracking the player with top surface 110x*. As depicted, the FOV's for cameras **140-b** and **140-d** are almost fully blocked by other players whereas the FOV for camera **140-a** is partially blocked but the FOV for camera **104-c** is clear. *The preferred embodiment will automatically reassign cameras such as 141-d that may already be tracking another player, (e.g., the player with top surface 111x) to now follow a different player with top surface 110x so as to ensure total maximum player visibility.* This reassignment decision can be based upon the information gathered by the scalable area tracking matrix **504m**, predictive calculations made by computer **160** concerning the expected next positions of any and all players, or

both." (italics added).

c. Clearly, my teachings are addressing problems beyond the scope and capabilities of Paff's system, such as solving for the best view of one occluded object in a group of multiple objects as videoed from any of multiple viewing angles.

11. "for controllably detecting the motion of each object in (X, Y, Z) space (fig. 6, slave cameras, col. 4, lines 50 – 60)"

a. Paff's master and slave cameras only provide the current pan and tilt angles of their optical axis to his controller 10. Based upon his technique of essentially projecting the optical axis from the master camera through an assumed object height (e.g. 4 feet) off the premises floor, Paff provides *at most* only a single point of data concerning each object at any given instant. This single point of data is assumed to be centered on a human body at the assumed height.

i. In the discussion of my Fig. 16b, my specification states:

"As depicted in Fig. **26b**, at least one point on sticker **109tm&id** that is in view of both fixed pre-calibrated camera **124** and movable camera **140** is first located in local rink (X, Y) coordinates based upon information provided by camera **124**. Once located, the same point is analyzed by computer **160** from the images captured by camera **140** along with other measurable information such as the current rotations of the panning and tilting mechanisms supporting camera **140** as well as the zooming mechanism associated with its lens.

*** During analysis, the determined (X, Y) location of the captured point *[on the object]* is used to center the (X, Y, Z) coordinate system of *[second or slave]* camera **140**. Once centered, the (Z) height scale can be set and then used to apply to all other common points in view of both the (X, Y) *[first]* camera **124** and the (X, Y, Z) *[second or slave]* camera **140**. *These points include not only those on helmet sticker **109tm&id** but also those **throughout all the body of player 17**.*" ***

(italics, bolding, brackets and *** added)

1. The ability to gather additional feature information about each tracked object has critical value to my teachings as it allows the objects to be represented as a mathematical model of their outer structure (e.g. body-points on a human.) This mathematical model can then be transmitted at minimal bandwidths and later reassembled into a graphic representation of any tracked objects. Obviously, a single (X, Y, Z = 4 ft.) point on a given object is not sufficient to recreate a realistic rendition of the actual movement of an athlete during a sporting event. (See my Fig's 22c and 22d, as well as their descriptions.)

12. "and a second algorithm operated on the computer system responsive to the second set of cameras"

a. By definition, at any instant during usage, Paff's system has only one master camera and potentially multiple slave cameras. While Paff does teach that any of his slave cameras can be switched with the master camera, his invention still only outputs a single (X, Y) stream regardless. Furthermore, this stream is always coming from the current master camera, and never from the current slave cameras.

    i. Hence, unlike my teachings, Paff's slave cameras do not gather information but rather only respond to information generated by his controller 10 in response to the master camera movements.

b. Therefore, Paff has no equivalent to my second algorithm that is responsive to the second set of cameras (i.e. his slave cameras.) Again, Paff's controller 10 and its mathematical algorithm at most approximates my first algorithm, and is never responsive to *image data* or in fact *any data* generated by a slave camera.

13. "for determining the (X, Y, Z) dimensional characteristics of each object,"

    a. As previously stated, Paff only continuously determines one (X, Y) point of information about at most a single object or single group of objects. This single points is the a "position in the premises 1 which can be selected by the master cameras MD." (see Paff's Fig.'s 1 – 3 depicting the X, Y grid of the premises.)

        i. This (X, Y) premises location is equivalent in function to the (X, Y) object location determined by my first algorithm operation on the first stream of video.

        ii. This (X, Y) premises location is not equivalent to the (X, Y, Z) three dimensional object feature location information created by the additional image analysis of the second stream of video.

    b. Also as previously stated, my teachings show how to find the (X, Y, Z) locations of multiple feature points on a single object (such as body points on an athlete.) These (X, Y, Z) dimensional characteristics referred to in my application are in addition to the general size of the object and the (X, Y) "premises" location of its assumed center. (See my Fig.'s 22c, 22d and 26b.)

    c. Paff at most can be construed to teach how to determine a *single* point of interest on a *single* object, whereas I teach determining *multiple* points of interest on *multiple* objects. This is a significantly harder task requiring substantially different apparatus and method steps.

14. "(fig. 6, each slave camera has the same component configuration as the master camera, each has a controller 10, which includes a mathematical processor to perform mathematical calculation based on the information received from the master camera)."

    a. While Paff's physical structure of the slave and master cameras is substantially identical, the behavioral "state" characteristics are significantly different between the master and slaves as is shown by the following excerpts from Paff's teaching.

        i. The operator selects one master camera. Paff Col. 7, lines 35-37:

"It should be noted that any of the cameras in the system 1 can be switched to take the role of *the* master camera. The operator at the station 11 can effect this change." (italics added).

ii. The operator then adjusts the one master camera to follow a single object (i.e. "the subject S".) Paff Col. 6, lines 4-14:

"To track the subject S, the operator at the control station 11 moves his or her joy stick control 11B to provide signals to the master camera MD. These signals cause the camera to adjust its pan and tilt angles to train the optical axis 17A of the image and lens assembly 17 of the master camera so that this axis intersects the coordinate position (36, 32) of the subject. Accordingly, the master camera now views the subject and the image of the subject appears on the monitor 11A. The master camera MD *[via its controller 10]* then calculates the coordinate position of the subject S in the plane 2." (italics added.)

iii. The controller for the master camera then makes a series of calculations to determine the assumed center of the single object (i.e. "the subject S".) Paff Col. 6;

Lines 18-20: "The master camera MD first calculates via its controller 10 the distance Dmd in the plane 2 between the master camera MD and the subject S..."

Lines 27-31: "Referring now to Fig. 3, with the distance Dmd known, the master camera MD then, via its controller 10, determines the coordinate position in the plane 2 of the subject S relative to the projected coordinate position of the master camera in the plane."

Lines 39-41: "The controller 10 of the master camera MD then performs three calculations based on the above information."

iv. The controller for the master camera then broadcasts the operator selected center-of-view position to the slave cameras. Paff Col. 6, Lines 56-58:

"The master camera MD then broadcasts the coordinate position of the subject S and a desired range value to all the slave cameras SD1-SD5."

v. Each slave camera, based upon its pre-known fixed location and its current pan and tilt orientation, then calculates its own adjustments to its pan and tilt angles so as to keep its optical axis focused on the object (i.e. subject S.) Paff Col. 6;

Line 62-65: "More particularly, the slave camera SD2 first determines the distance between its projected position ... and the received position ... of subject S."

Lines 66-68: "Thus, referring to Fig. 5, the slave camera SD2 determines the X and Y offsets between itself and the subject S…"

Col. 7, Lines 18-20: "The slave camera SD2, knowing the distance Dsd2 to the subject S and the offset X can calculate the adjusted pan angle…"

Col 7, Lines 27-28: "Referring to Fig. 4, the slave camera SD2 then calculates the tilt angle…"

vi. In recap response to this clause of the OA, while the apparatus of Paff's master and slave cameras may be identical, at any given moment there is only one master with potentially many slaves. Hence, behaviorally, there is a clear difference between master and slave that is important when considering a comparison with my teachings. The master is always and only used to determine the (X, Y) location of the single object (Subject S) and the slaves never determine any (X, Y) let alone (X, Y, Z) information regarding Subject S.

   1. Given this understanding, the only possible comparison between my teachings and that of Paff's must attempt to equate my:

      a. First set of fixed cameras to Paff's single master camera, and

      b. Second set of movable cameras to Paff's multiple salve cameras.

   2. Therefore, the slave cameras cannot be reasonable "combined" with the master camera to imply that Paff teaches a "first set of cameras" nor can the master's "independently calculating controller algorithm" be attributed to each slave camera (i.e. the slaves do not calculate independent center-of-views but rather all are centered on the (X, Y) premises location selected by with master.)

***(3) A side-by-side comparison of the changes to my claims:***

Attached please find my new Claims 50-92 replacing my herein cancelled Claims 34-49. As a matter of record, and in accordance to the arguments provided above, I respectfully submit that my original Claims 34-49 remained allowable as is, without the need for further revision in order to overcome Paff. All of the new Claims 50-92 are substantially identical in both apparatus and method to my Claims 34-49 and have only been revised to *increase* clarity and direct comparison with Paff.

Please note that cancelled apparatus Claims 34-43 have been replaced by new apparatus Claims 50-59. New Claims 60-70 recast new Claims 50-59 with wording and independent structure that is even more directly comparable to Paff. Similarly, cancelled method Claims 44-49 have been replaced by new method Claims 71-80. New Claims 81-92 recast new Claims 71-80 with wording and independent structure that is even more directly comparable to Paff.

More specifically, the following table provides a side-by-side comparison of the now cancelled independent apparatus Claim 34 with the new independent apparatus Claims 50 and 60.

| **Independent Apparatus Claims** | | |
|---|---|---|
| **Claim 34** | **Claim 50** | **Claim 60** |
| *(now cancelled)* | *directly parallel with Claim 34* | *restates Claim 50 even more comparable to Paff* |
| An automated system for tracking the movement of multiple objects within a predefined area comprising: | An automated system for tracking the movement of one or more objects within a predefined area based upon computer analysis of captured video images and not requiring operator intervention, comprising: | An automated system for controlling some combination of at least the pan, tilt and/or zoom controls of one or more second movable cameras as they video the activities of one or more objects in a predefined area, where the control signals directing the second movable cameras are automatically generated without operator intervention and based upon computer analysis of video images captured by one or more first fixed cameras that together form a contiguous and continuous view of the same area, comprising: |
| a first set of cameras forming a fixed area tracking matrix for first detecting the motion of each object in (X, Y) space; | a first set of fixed cameras for generating a first video stream of images that together form a contiguous and continuous view of the predefined area; | (same as Claim 50) |
| a first algorithm operated on a computer system responsive to the first set of cameras for determining the (X, Y) location of each object | a first algorithm operated on a computer system responsive to the first stream of video images for analyzing those images to first determine the relative (X, Y) location and dimensional characteristics of each object within the predefined area and for forming a tracking database representative of each object's locations, movements and dimensional characteristics; | (same as Claim 50) |
| a second set of cameras forming a movable volume tracking matrix responsive to the determined (X, Y) locations from the first algorithm for controllably | a second set of movable cameras responsive to the tracking database, wherein each movable camera is automatically directed without operator intervention to | (same as Claim 50) |

| | | |
|---|---|---|
| detecting the motion of each object in (X, Y, Z) space; and | maintain an independent view of one or more objects within the predefined area and where the second set of movable cameras continuously outputs a second stream of video images, and | |
| a second algorithm operated on the computer system responsive to the second set of cameras for determining the (X, Y, Z) dimensional characteristics of each object and for forming a database representative of each object's locations, movements and dimensional characteristics. | a second algorithm operated on a computer system responsive to the second stream of video images for determining additional relative (X, Y, Z) location and dimensional characteristics of each object and for updating the tracking database. | (added as 2nd dependent claim) |

The following table provides a side-by-side comparison of the new independent method Claims 71 and 81 with respect to the new apparatus Claim 50.

| Independent Method Claims | | |
|---|---|---|
| **Apparatus Claim 50** | **Claim 71** | **Claim 81** |
| *(used as a basis instead of cancelled method Claim 44)* | *directly parallel with new Claim 50* | *restates Claim 71 even more comparable to Paff* |
| An automated system for tracking the movement of one or more objects within a predefined area based upon computer analysis of captured video images and not requiring operator intervention, comprising: | A method for tracking the movement of one or more objects within a predefined area based upon computer analysis of captured video images and not requiring operator intervention, comprising the steps of: | A method for controlling some combination of at least the pan, tilt and/or zoom controls of one or more second movable cameras as they video the activities of one or more objects in a predefined area, where the control signals directing the second movable cameras are automatically generated without operator intervention and based upon computer analysis of video images captured by one or more first fixed cameras that together form a contiguous and continuous view of the same area, comprising the steps of: |
| a first set of fixed cameras for generating a first video stream of images that together form a contiguous and continuous view of the predefined area; | capturing a continuous first stream of video images using a first set of fixed cameras, wherein the images together form a contiguous and continuous view of the predefined area; | (same as Claim 71) |
| a first algorithm operated on a computer system responsive to the first stream of video images for analyzing those images to first determine the relative (X, Y) location and dimensional characteristics of each object within the predefined area and for forming a tracking database representative of each object's locations, movements and dimensional characteristics; | detecting the (X, Y) location and dimensional information of each object relative to the predefined area using computer based image analysis of the first stream of video images without the aid of an operator; | (same as Claim 71) |
| a second set of movable cameras responsive to the tracking database, wherein each movable camera is | using the detected (X, Y) location and dimensional information regarding each object to automatically and | (same as Claim 71) |

| | | |
|---|---|---|
| automatically directed without operator intervention to maintain an independent view of one or more objects within the predefined area and where the second set of movable cameras continuously outputs a second stream of video images, and | individually direct some combination of at least the pan, tilt and/or zoom movements of a each camera in a second set of one or more movable cameras; | |
| | capturing a continuous second stream of video images using the second set of automatically movable cameras, wherein the images create independent views of one or more of the objects within the predefined area; | (same as Claim 71) |
| a second algorithm operated on a computer system responsive to the second stream of video images for determining additional relative (X, Y, Z) location and dimensional characteristics of each object and for updating the tracking database. | detecting additional (X, Y, Z) location and dimensional information of each object viewed using computer based image analysis of the second stream of video images without the aid of an operator, and | (added as 2$^{nd}$ dependent claim) |
| | combining the information detected by image analysis of both the first and second video streams into a continuously updated tracking database indicating the relative (X, Y, Z) locations and dimensional information of the objects relative to the predefined area. | (added as 2$^{nd}$ dependent claim) |

*(3) Arguments for the allowance of the revised claims:*

Based upon the above provided analysis, I respectfully request that you allow my new claims as revised and included on the attached document especially for the following reasons:

1. I teach and claim:
    a. "An automated system," that does "not require operator intervention":
        i. Paff's system cannot function without operator intervention and is therefore not an truly automated system.
    b. A system "for tracking the movement of one or more objects":
        i. Paff's system does not track objects in a literal and important sense but rather tracks the single optical axis of a Master Camera, representing a single "selected position" as chosen by an operator;
            1. hence, there is no assurance that the single optical axis is always pointed at the single object to be tracked let alone that objects meaningful center-point;

        a. this is especially true for objects experiencing high acceleration / deceleration that will be harder for a human operator to manually follow;
    2. when tracking non-human objects, such as a hockey puck, basketball or football, their comparatively small size and large variation in distance off the ground as compared to a human precludes Paff's simplified assumptions for calculating current (X, Y) coordinates based upon a fixed object height (e.g. 4 ft) algorithm, and
    3. Paff only ever speaks of a single stream of (X, Y) coordinates which will not sufficiently describe the random, uncorrelated movement of multiple objects.

c. A system "for tracking the movement" … "within a predefined area":
    i. Paff's system is limited to only tracking objects that an operator can physically see looking through a Master Camera and multiple Slave cameras;
        1. All of these cameras are specified as movable cameras and therefore there is no structural guarantee that all of the predefined area will be in view of the operator at all times, and
        2. Since Paff specifies that the Slave cameras should always adjust their center-of-view to match that of the master camera, this even further ensures that the Master and Slaves together are always excluding at least some portion of the predefined area;
            a. And any excluded portion of the predefined area could experience the presence of an object to be tracked, which Paff would miss.

d. A system that bases its tracking calculations on "computer analysis of captured video images":
    i. Paff's calculations of the (X, Y) coordinates of the Master Camera's currently "selected position" are primarily based upon control signals representing at least the current pan and tilt angles of the Master Camera.
        1. Paff never teaches the need or use of analyzing captured video images.

e. A system that employs a "first set of fixed cameras":
    i. Paff does not specify the use of any fixed cameras, at all, and therefore cannot also guarantee a continuous and contiguous full-view of the predefined area.

f. A system that "generates a first video stream of images that together form a contiguous and continuous view of the predefined area":
    i. Paff's first video stream is from the single master camera and it does not provide a contiguous view of the predefined area.

g. A "first algorithm" … "responsive to the first video stream":
    i. Paff's first algorithm, his controller, is responsive to the stream of control signals generated by the joy stick movements. No where does Paff discuss using image analysis on the video generated by the Master Camera.

h. A system capable of "determining dimensional characteristics" of each tracked object:
    i. Paff only teaches the creation of a single stream of (X, Y) "premises" coordinates; again, representative of the currently "selected position" and not even of the object, let alone multiple objects.

        1. Paff does not show any way of determining the size characteristics of a tracked object, which would necessarily require some form of image analysis on a stream of video.

i. A system with a "second algorithm" … "responsive to the second stream of video images for determining relative (X, Y, Z) location and dimensional characteristics of each object."
    i. Paff does not teach any second algorithm that operates on the images that are being collected by his slave cameras, let alone an image analysis algorithm similar to my teachings for determining (X, Y, Z) object feature locations;
    ii. Furthermore, one of the major goals of my patent is to create an efficient system for collecting maximized perspective views of free-moving multiple objects (e.g. such as players in a sporting contest) so that the continuous stream of images from these perspective views could themselves be analyzed providing significant additional (X, Y, Z) feature information not easily obtained from fixed cameras alone;
        1. This goal needed to be accomplished automatically, without requiring an operator to direct the collection of these perspective views. This in turn requires the first set of fixed cameras to provide the basis for object tracking through image analysis where the tracked movements provide the data to optimize the perspective views of the continuously evolving physical configuration of the multiple objects within the predefined area.
        2. In contrast, Paff teaches no additional value from analyzing the video collected by the Slave cameras. Precisely because this is not a goal of his system and also therefore not a sufficiently optimized possibility due to his design.

These key elements, features and functions of my teachings are not claimed, specified or anticipated by Paff. Therefore, I respectfully request the allowance of my new claims 50 – 92.

I thank you for your consideration in these matters.

Sincerely,

James A. Aman

This communication was sent by U.S. Postal Service, "Express Mail Post Office to Addressee" on 6/12/06. The Express Mail Label number is ED669689731US.